NEW APPROACH TO STOCHASTIC MOMENTUM COOLING

W. Kells

January 1980

I. Introduction

The Fermilab $\bar{p}$ source group has recently proposed a design for a high-luminosity $\bar{p}$ source.[1] To utilize electron cooling for $\bar{p}$ stacking in this design a large longitudinal phase space compression was required to match a maximum feasible space at production energy (340 ev-s @ 4.5 GeV) to the existing electron ring (2.7 ev-s @ 625 MeV). The candidate process for this compression, filter stochastic cooling,[2] can only cool by modest factors (<10×) in a few seconds. Figure 1 summarizes current performance of stochastic-cooling systems.

The solution to this dilemma was to cool by small factors (<10×) in stages.[1,3] Between each stage the cooling system's parameters are readjusted by deceleration such as to restore the initial conditions of fast cooling. A factor of 125 compression of longitudinal phase space in 6s of cooling was conservatively predicted. Since $\bar{p}$ deceleration (production energy to electron "freezer" energy) had always been a step in collection, such a stochastic cooling technique was a natural one to adopt. In contrast, the CERN AA project requires no deceleration and consequently no large fast initial cooling step.

Since each stage of cooling is to be identical to the, by now well-understood CERN approach, the Fermilab group was able to make their proposal with great confidence given the small number of stages involved.[3,4] In this paper the idea of combining stochastic cooling with deceleration is developed to its logical (and therefore some what more speculative) limits. Some of the conclusions are:

1. The overall cooling "rate" ($F/\Delta T$) is optimized by cooling in the limit of an infinite number of cooling stages;
2. In this limit filter cooling is not necessarily the technique of choice;
3. A more natural technique is proposed;
4. This new technique allows a far simpler beam pick-up to be used;
5. An increased useful factor of transverse cooling is allowed "for free."

II. Comparison of Filter Cooling With and Without Deceleration

Throughout this paper I will only **compare** cooling schemes and therefore shall not attempt to calculate absolute cooling rates or times. The full equations describing stochastic cooling are nonlinear.[4] I therefore avoid any numerical solutions. Although simplifications of the exact equations will arise I consistently apply them in such a way as to only emphasize the comparison sought.

The model I use for filter momentum cooling is

$$\dot{\psi} = \frac{\partial}{\partial \varepsilon} \left(g\varepsilon\psi - g^2 \kappa \varepsilon^2 \psi \frac{\partial \psi}{\partial \varepsilon} \right)$$

$$\psi(\varepsilon, t) = dN/d\varepsilon \qquad (1)$$

$$\varepsilon = E - E_0 ,$$

where E_0 is the nominal cooling energy. This equation is identical to the CERN Fokker-Planck description except for some simplifying assumptions about the filter channel characteristics:

1. The electronic channel gain is a constant throughout its bandwidth W. In particular g and κ then become constants.
2. All particles are assumed to be within a linear region of the filter response. This means $1/2W \geq (\sigma_E/E_0)\beta\eta T_0$.
3. The filter and electronic channels are lossless and noiseless. Loss may be represented by a dependence ε^2 + const.) in the second term instead of ε^2.[4,6]

Under these conditions the coefficients g and κ may be evaluated. In Appendix I, I derive (1) and hence κ starting from a time domain point of view. It is shown that if this approach is handled with sufficient care[5,6] that after transformation to frequency domain the result is identical to the CERN derivations.

Already (1) may be used to illustrate the advantage of decelerating [i.e., decreasing $\eta \equiv (d\varepsilon/\varepsilon_0)/(d\omega/\omega_0)$] during the cooling process. The quantity of interest in longitudinal cooling is $\sigma_\varepsilon \equiv (\int\varepsilon^2\psi d\varepsilon)^{1/2}$ where $\int\psi E \equiv 1$. Equation (1) may be integrated to give $\dot{\sigma}_E$:

$$R_\varepsilon \equiv \frac{\dot{\sigma}_\varepsilon}{\sigma_\varepsilon} = -g + 3/2\ g^2\kappa\ \{(\int\varepsilon^2\psi^2 d\varepsilon)/\sigma_\varepsilon{}^2\} \tag{2}$$

which implies a maximum instantaneous rate given κ and the instantaneous distribution $\psi(\varepsilon,t)$:

$$g_{max} = \frac{1}{3\kappa} \; (\int \varepsilon^2 \psi^2 d\varepsilon)/\sigma_\varepsilon{}^2\}^{-1}$$
$$\Rightarrow \frac{\dot{\sigma}}{\sigma} \Big|_{max} = -\frac{1}{2} g_{max}. \tag{3}$$

For fast cooling schemes, where we seek maximal rates, the cooling rate must continuously decrease as σ_ε decreases since the ratio $\{(\int\varepsilon^2\psi^2 d\varepsilon)/g_\varepsilon{}^2\}^{-1}$ continuously decreases. Let $F[\equiv \sigma_\varepsilon(\text{final}) / \sigma_\varepsilon(\text{initial})]$ be the net cooling factor achieved in ΔT. The best one can achieve with fixed κ (fixed E_0) is to vary g during the cooling such as to satisfy (3). The appropriate description of this process is the function $\Delta T[F, g_{max}(t)]$ which, since (1) is nonliner, is obtainable only via numerical solution.

If we decelerate in stages during the cooling (in a manner detailed in Ref. 1), then κ is no longer a constant. It is proportional to η (see Appendix I), thus allowing the growth of $(\sigma_\varepsilon{}^2/\ \varepsilon^2\psi^2 d\varepsilon)^{-1}$ to be compensated for. Let us neglect for the moment (until section IV) the time required to (adiabatically) capture, decelerate, and debunch the beam. Then in the limit of an infinite number of cooling/deceleration stages the maximal cooling rate would be maintained at its initial value:

$$\frac{\dot{\sigma}_\varepsilon}{\sigma_\varepsilon}\Big|_{max} = \frac{1}{6\kappa(t=0)} \left\{(\int\varepsilon^2\psi^2_{t=0}\; d\varepsilon)/\sigma^2{}_{\varepsilon,t=0}\right\}^{-1}$$
$$\Rightarrow \Delta T(F,\ \kappa) = \left(\frac{\dot{\sigma}_\varepsilon}{\sigma_\varepsilon}\Big|_{max}\right)^{-1} \text{LnF}. \tag{4}$$

In Appendix II I derive a lower bound for $\Delta T(F, g_{max})$ (the fixed E_0 time evolution):

$$\Delta T(F,\ g_m(t)) \geqslant \left(\frac{\dot{\sigma}_\varepsilon}{\sigma_\varepsilon}\Big|_{max}\right)^{-1} (F-1) \tag{5}$$

which we compare to (4) in Fig. 2. The comparison also illustrates the even lower performance situation of fixed gain filter cooling $[\Delta T(F,\ g = \text{const.})]$. The lower bound (5) corresponds to g decreasing in time as:

$$g(t) = g(0)\sigma_\varepsilon(t)/\sigma_\varepsilon(0). \qquad (6)$$

For the Fermilab precooler, where $F \geqslant 125$[1] the implication of the above comparison is illustrated in Fig. 3.

The integrated form (2) of the Fokker-Planck equation differs qualitatively from elementary deviations of the rate from time domain averaging of turn-to-turn corrections.[7] Confusion has existed in reconciling the frequency domain and these elementary time domain rates. Sacherer, for instance, has "derived" this naive time domain result starting from (1).[8] Beyond linearizing the kick in his derivation, Sacherer was forced to linearize the ψ dependence itself. These original naive rate equations are too drastic a simplification of the physics. It is imperative to retain the ψ^2 dependence as in (2).

III. A New Momentum Cooling Method.

If we choose continuous deceleration (CD) cooling then we must re-examine whether filter feedback is the highest performance technique. Keep in mind that the filter technique was invented as an improvement over the **fixed** momentum Palmer method.[9] The advantages were:

1. Allows the use of sum pickups of uniform (transversely) high sensitivity.
2. Since the filter comes after initial low noise preamplification, electronic noise is suppressed near the filter gain zeros. Since the particles condense progressively closer to the zeros, the relevent part of the electronic noise spectrum[10] is also amplified progressively less. If electronic noise were the only diffusion term source then, since its spectrum is flat and constant, we could cool indefinitely (with a perfect filter and at an ever diminishing rate) with **constant** electronic channel gain.
3. Schottky noise is suppressed by the same mechanism as in 2. However this diminution of diffusion is more than offset by the central growth in particle density.

But the filter cooling is by no means ideal. The growth of $\int \varepsilon^2 \psi^2 / \sigma_\varepsilon{}^2$ determines a falling $g_{max}(t)$ even for a case of continuously optimized g. It is just CD cooling which remedies this, for any cooling technique.

There is another inherent disadvantage of filter cooling. Implicit in correct use of the Fokker-Planck approach is an averaging of the diffusion term over a time τ long compared to the pair correlation time τ_2.[11] This means that the diffusion term in (1) must be proportional to a "mixing factor," i.e., number of turns it takes to separate all initial particle pairings by a (longitudinal) distance equal to the effective pickup length ($\approx c/2\ W_{eff}$, W_{eff} = effective system bandwidth). Now, we can always have a mixing factor = 1 (<=> "good mixing") by choosing η large enough. On the other hand, filter cooling depends upon a turn-to-turn correlation within the effective pickup width for its basic operation. In practice, mixing factors ≥ 5 are apparently necessary.[12] Thus the diffusion term is ≥ 5 times ideal; the maximal cooling rate then being $\leq$ 1/5 ideal [Eq. (3)].

I propose a new momentum-cooling technique which retains features 1 and 2 while allowing a mixing factor = 1. On the other hand, the advantage 3 is lost which I will argue leaves a net gain in performance.

This new method is schematicized in Fig. 4b. The "signal" differentiating particle's momenta is derived from PU-KICKER transit time. Therefore I shall refer to it as the "transit time" ("TT" contrasted to "FT" for filter technique) method. In principle the parameter determining the coherent signal correction ($\equiv \eta^*$) is independent of the mixing parameter η for TT cooling ($=\eta'$) whereas only one parameter is available to optimize FT cooling ($\eta^* = \eta$).

Actually the two methods are quite similar. Let $\mathcal{E}_0 h'(t)$ be the axial electric field impluse produced at the kicker center by one particle (same for all particles). Let the axial electric field experienced by a particle (velocity = c) transiting the kicker with a constant potential applied to the kicker be proportional to f(t). The single particle self correction is the convolution of these functions

$$\delta\varepsilon(\tau_{TT}) = e\,\mathcal{E}_0 \int_{-\infty}^{\infty} h'(t)\, f(t-\tau_{TT})\, dt$$
$$\tau_{TT} \equiv \varepsilon\beta\eta'\tau/E_0 . \qquad (7)$$

The ideal filter cooling circuit is identical to the one just described for TT cooling with the addition of a stub line and with $\eta'_{FT} = 0$ (Fig. 4a). Therefore its single particle self correction is

$$\delta\varepsilon(\tau) = \frac{1}{\sqrt{2}}\, e\,\mathcal{E}_0 \int_{-\infty}^{\infty} \{h'(t) - h'(t-\tau)\}\, f(t)dt$$
$$\tau = \varepsilon\beta\eta T/E_0 \qquad (8)$$

But for $\eta'_{TT} = \eta_{FT}$ (7) and (8) are exactly proportional since the symmetry of the response dictates that $\int h' f dt = 0$. This shows that these two cooling methods have basically identical friction terms.

On the other hand, the diffusion terms are different. The influence of the jth particle on the ith is

$$\text{TT:} \quad \delta\varepsilon_{ij}(\tau) = e\mathcal{E}_0 \int h'(t+\Delta t_{ij})\, f(t-\tau_j)\, dt \qquad (9)$$

$$\text{FT:} \quad \delta\varepsilon_{ij}(\tau) = \frac{1}{\sqrt{2}}\, e\mathcal{E}_0 \int \{h'(t) - h'(t-\tau_j)\}\, f(t+\Delta t_{ij})\, dt \qquad (10)$$

The second vanishes identically as τ_j ($\propto \varepsilon_j$) $\rightarrow 0$. The diffusion term is the average of the $\delta\varepsilon^2_{ij}$. One is lead to a a quadratic ε_i dependence for FT whereas the TT average is i independent.

Maximum cooling rate is determined by the magnitude of the diffusion turn. The proper comparison of FT rate to TT rate is the initial rate with $\eta'_{TT} = \eta_{FT}$, since the FT rate is maximum initially. Appendix I calculates these diffusion coefficients [the κ of Eq. (1) and its equivalent for TT cooling] for an initial distribution all the particles of which lie within a linearizable region of the response. The result is

$$\frac{\kappa_{TT}}{\kappa_{FT}} = \frac{1}{2} \times \frac{5}{2} = \frac{5}{4}, \tag{11}$$

where the 5/2 comes from the averages and κ_{TT} is defined in (17). The comparison of initial rates [see Eq. (2) and its equivalent for TT cooling derived from (16)]:

$$\frac{R_{TT}}{R_{FT}} = \frac{3}{4}\left(\frac{K_{TT}MF_{TT}}{K_{FT}MF_{FT}}\right)^{-1} \left(\int_{-\varepsilon^*}^{\varepsilon^*}(\varepsilon/\varepsilon^*)^2 \ \psi^2{}_0\, d\varepsilon \Big/ \int_{-\varepsilon^*}^{\varepsilon^*}\psi_0{}^2\, d\varepsilon\right), \tag{12}$$

where MF $\equiv$ mixing factor.

The parameter ε^* is the edge point of linear response in the feedback channel; approximately $\varepsilon^* \approx (4W\beta\eta' T_0/E_0)^{-1}$. To minimize the effect of electronic noise we must have the distribution maximally fill this response window. Choosing a square distribution ψ_0 then, with edges at $\pm\varepsilon^*$ gives (see Fig. 5)

$$\int(\varepsilon/\varepsilon^*)^2\psi_0{}^2 \ dE/\int \ \psi_0{}^2 dE = \frac{1}{3} \tag{13}$$

and

$$\frac{R_{TT}}{R_{FT}} = \frac{1}{5}\,\frac{MF_{FT}}{MF_{TT}}. \tag{14}$$

As noted above $MF_{FT} = 5$ seems minimal so that $R_{TT}/R_{FT} = 1$ can be attained simply by having the PU-K separation in TT cooling be 1/5 the cooler circumference (recall that $\eta'_{TT} = \eta_{FT}$ for this comparison).

In fact the above comparison **favors** FT cooling in at least two respects. First, the deleterious effect of PU → K dispersion is neglected ($\eta'_{FT} = 0$ is assumed). Second, filter losses and imperfections (e.g., non-multiple notch minima) are neglected. In fact, the practical advantage of **requiring no notch filter** may be the strongest point in favor of the TT cooling. The differentiation necessary could be a shorted stub, but its length is $\ll c/4W$.

In principle MF_{TT} can be made <1 by having sufficiently high η and having multiple cooling circuits, so that two cycles of momentum correction and mixing occur each turn. Because it requires turn-turn frequency correlation, no such improvement is possible with FT cooling.

One may ask whether (14) is reasonable. An exact answer can be had only from examining the shape evolution of the density. For instance, a Gaussian profile which is 95% (2σ) inside the linearized response bounds gives the value of $1/4\sqrt{2}$. Evidently a value somewhere between 1/3 and 1/6 is reasonable.

IV. Mode of Deceleration

It is important to realize that the above analysis of TT cooling is for the limit of an infinite number of fixed-energy cooling steps. One must imagine an instantaneous deceleration of the beam between each of these steps. In practice, one must instead, smoothly decelerate the beam while cooling. The very important problem of stochastically cooling rf bunched beams I shall not directly analyze. In the last section of this paper I comment some more on this. Implicit at all other places is the assumption that a mode of rf capture can be devised that will not degrade the cooling effect.

The problem of cooling bunched beams is not the only issue brought up with smooth deceleration. When a beam is decelerated the rotation frequency scale shrinks in proportion to η (the lattice is at all times below transition). In other words, deceleration induces a pseudo-force into the time evolution of ψ. The sign of this force is as an anti-friction. Presumably the initial distribution in frequency space maximally fills the feedback circuit response bandwith (Fig. 5). Therefore the pseudo-force must be at least balanced by the cooling term or the beam will be lost from the notch. (This is very similar to the increase in rf voltage needed for capture with deceleration rate.) Since this crucial stability condition is manifest in frequency space, we write the Fokker-Planck equation and the distribution as functions of $\omega = 2\pi f_0 \eta \varepsilon / E_0$. The effect (pseudo-force) of deceleration alone is

$$\dot{\psi}\ (\omega) = \frac{\dot{\omega}}{\omega} \frac{\partial}{\partial \omega} (\omega\psi), \tag{15}$$

where

$$\frac{\dot{\omega}}{\omega} = -\dot{E}_0 \frac{d}{dE} \ell n \left(\frac{n}{E_0}\right) \text{ (fixed E and } f_0\text{)}.$$

The Fokker-Planck equation for TT cooling (see Appendix I) is

$$\begin{aligned} \dot{\psi}\ (\omega) &= \frac{\partial}{\partial \omega} \{g\omega\psi - g^2 \kappa_{TT}\psi \frac{\partial \psi}{\partial \omega}\} \\ &\quad + \text{ pseudo force} \\ &= \frac{\partial}{\partial \omega} \{(g-s)\omega - g^2 \kappa_{TT}\psi\psi'\} \\ &\quad \text{with } S \equiv \dot{\omega}/\omega. \end{aligned} \tag{16}$$

The criteria for fixing the deceleration rate $\dot{E}_0$ will then be to maximize g and s with the constraint that the second moment of (16), $\sigma_\omega/\sigma_\omega$, = 0. The result is

$$S_{max} = (2\ \kappa_{TT})^{-1} \ (\int \psi^2 dE/\sigma_E{}^2)^{-1}. \tag{17}$$

This is the same value that (16) would have yielded for $\dot{\sigma}_\omega/\sigma_\omega|_{max} = \dot{\sigma}_\varepsilon/\sigma_\varepsilon$ at **fixed** energy (i.e., **no** pseudo-force). Consistent with the approximations (**second moment rates**) we have been using, the above deceleration criteria gives

$$\dot{\sigma}_\varepsilon/\sigma_\varepsilon = 3/2\ S_{max}. \tag{18}$$

Thus a factor 3/2 difference exists between the infinite limit of step cooling and smooth deceleration. Notice that the 3/2 factor **favors** the smooth deceleration, changing Eq. (14) to

$$\frac{R_{TT}}{R_{FT}} = \frac{3}{10}\,\frac{MF_{FT}}{MF_{TT}}. \qquad (14b)$$

V. Corollaries to Transit Time Cooling

Although maximum cooling rates are proportional to the effective bandwidth (W_{eff}) of the feedback system, practical systems[13] are limited to $W_{eff} \lesssim$ 200-300 MHz for three reasons:

1. Poor response time PU and kicker cores ($2\ W_{eff} \sim 1/\tau$) aggravated by the necessity of many closely packed cores for signal/noise enhancement.
2. Low level electronic bandwidth limitations. **Percentage** flat bandwidth is the relevant criteria for (e.g.) amplifiers. 50-250 MHz represents a **500%** bandwidth; this is considered very broadband. Noise figure and bandwidth trade off.
3. Power Amplifiers. C. W. power and **percentage** bandwidth also trade off.

The usual communications engineering answer to demand for larger W_{eff} is "go to higher frequency." That is, the **percentage** bandwidth requirement relaxes. In a recent note I describe a waveguide cumulative pick up (particle velocity matched to phase velocity) which physically **must** operate at microwave frequency (4-5 GHz).[14] There is still a bandwidth limitation imposed by the impossibility of matching phase and group velocities in such a structure; however, the waveguide is enormously simpler than hundreds of ganged cores.

But there is a fundamental reason that higher center frequencies **cannot** be used with filter cooling. The harmonic numbers operative in the PU signal would be so large that Schottky bands would overlap[4] unless the ring η were unrealistically low. With transit time cooling we have seen that η can (should) be arbitrarily large; there is no restriction on $W_{eff}/f_0\eta$. In other words, the TT method does not depend on the cyclical nature of the machine.

Unfortunately there is as analogous problem if a microwave center frequency is carried on through to the kicker. This is illustrated in Fig. 5b. The problem is that the individual particle response (the function h(t) of the last section) is now **mot monotonic.** The analysis in Appendix I assumes this for the results on cooling previously discussed. Only a narrow band of particles which traverse the kicker within the **central** one-half cycle of Fig. 5b will be properly cooled. What we need is rectification just before the kicker (then a gap, see Fig. 6). Some advantages of such a microwave circuit are:

1. Elimination of complex **multiple** PU and K structures.
2. ~10% bandwidth for W_{eff} = 500 MHz @ 5 GHz.
3. In the microwave regime it becomes possible to use GaAs Fet preamps which have the lowest noise figures of any broadband device. Their best performance is in the few GHz region. They can be cooled (LN_2) for significant NF improvement.
4. TWT power amplifiers could be used. Hopefully power levels can be achieved that eliminate the need for multiple kicker arrays.

An exact design for rectifying the high power rf remains an unknown.

The Fermilab precooler design[1] has a momentum spread limited aperture at the $\bar{p}$ injection energy, but at 200 MeV (freezer ring energy) it is limited by vertical betatron amplitude. It is therefore natural to consider ways of performing betatron cooling on the scale of a few seconds.[15] Of course, accepting more transverse phase space from the $\bar{p}$ target (along with the same $\Delta p/p$ won't give an enhanced $\bar{p}$ accumulation rate since any stochastic (non-noise limited) cooling has its rate scaling as $(N_{total})^{-1}$. But betatron cooling would make the Fermilab scheme much more flexible; for instance, lower than expected target yields could be made up for.

In the case of continuous deceleration, a certain rate of transverse cooling (vertical in particular but also horizontal in the zero dispersion straight sections) can be accomplished which does **not** suffer from one of the two technical problems causing it to be slow (compared to momentum cooling). At fixed energy the sensitivity of a transverse PU is proportional to the ratio of beam size to PU aperture. The sensitivity diminishes as cooling proceeds. Actually the equivalent occurrs in FT cooling as the distribution shrinks toward the notch zero but the signal-to-noise ratio is typically worse for transverse PU's so that this point is crucial. However, if the rate of transverse cooling is just such that (**with deceleration now**) the cooling balances the deceleration induced blow up, then the aspect ratio does not change. The increase in allowed initial vertical acceptance is then equal to the factor of deceleration (~ factor of 9 for the precooler).

Since the mixing factor does not change during the deceleration cycle, such a factor of transverse cooling is then much more suited to the larger mixing required for maximal TT cooling. Fixed energy FT cooling is totally incompatable with simultaneous transverse coding since the MF increases as $(\Delta p/p)^{-1}$.

VI. Comments on Cooling of Bunched Beams

Naively, it seems that full rf capturing a beam would be fatal to momentum cooling since all particles would then have the same average rotation frequency (subharmonic of the rf frequency). However, mixing does occur because the synchrotron period is not constant at all points in the bucket and cooling is observed.[16] Especially encouraging is that the initial cooling rate is apparently undiminished from the case of no rf if the bunching factor is small. This suits the situation of TT cooling where η is continually increased to balance the momentum cooling.

This demonstration of bunched beam cooling also obviated an outstanding hardware question. That is, whether higher harmonics of the rf fundamental would inevitably leak into the supersensitive PU channel. As long as the rf harmonic is low (e.g., 2 for the Fermilab precooler) and the bunching factor small, spurious rf harmonics quickly die out (by h = 50 at ICE with rf h = 1).[16]

Since the coherent cooling signal for FT cooling depends upon a correlation in **instantaneous** (turn-to-turn) period the effect of rf on $\psi(\varepsilon)$ evolution for any ε is difficult to analyze. For TT cooling the effect is clear as long as the rf cavity is **not** inside the PU $\rightarrow$ K leg. Then there is **no** effect on the friction term. One must only worry about a possible diminution in global mixing. It has been pointed out that the normal **sinusoidal** rf waveform is merely a convenience and it is possible to choose a waveform (e.g., triangular) which would enhance mixing via a momentum dependence of synchrotron period.[17]

APPENDIX I

The form of Fokker-Planck equation (1) applicable to momentum cooling has been discussed in two places.[4,18] Neither analysis relates change in $\psi(\varepsilon,t)$ to the dynamics of the individual particles so that evaluation of the F and D coefficients is not very transparent. In particular, F and D are expressed as integrals in frequency (response spectrum of the circuit). Elementary derivations of the distribution rms evolution by averaging over single particle dynamics (therefore **time** domain) lead to:[9]

$$\dot{\sigma} \propto \frac{1}{T_0} \{-g + \frac{1}{2} \kappa g^2\} \qquad \text{(A1)}$$

a form clearly not derivable from a **nonlinear** equation in ψ [as (1)].

If, as I now show, the averaging leading to the second term in (22) is done correctly, one obtains expressions like (2) and a dynamically constructed expression for κ.

The calculation proceeds in the spirit of the F and D coefficient derivation by Ichimaru.[11] The key element is the existence of correlation time $\tau_2 \gg T_0$ ($\equiv$ revolution period which, for one feedback system, is the "collision time") which is the minimum time over which averages must be performed (i.e., $M = \tau_2/T_0$ revolutions). A given particle is influenced only by a subset of $N_s = N\, \tau_{PU}/T_0$ ($\tau_{PU} \approx 1/2\, W_{eff}$) of the total number N. On the average, its closest neighbor within N_s is spaced away in revolution frequency by $\overline{df} = (f_0 \eta/E_0)(\sigma_\varepsilon/N_s)$.

Therefore $\tau_2 \approx \tau_{PU}\,(\overline{df}/f_0)^{-1}$. For cooling to be considered as a Markov process, the correction must remain small on the τ_2 time scale. That is, (cooling rate)$^{-1} >> \tau_2$. For typical parameters being considered for fast precoolers ($2W_{eff}\,T_0 \approx 10^3$, $N = 5\times10^7$, $\approx$ "good mixing") we have $M \approx 10^4$ or $\tau_2 \approx 10$ ms which is small compared to maximal cooling rates expected of $\lesssim$ (100 ms)$^{-1}$. The Fokker-Planck approach is thus justified.

Using expressions (10-11) we can write the exact correction to the ith particle's momentum over the interval τ_2:

$$\varepsilon_i(\tau_2) = \varepsilon_i(0) - g \sum_{r=0}^{M} \sum_{j=1}^{N} \delta_{ijr}, \qquad \text{(A2)}$$

where δ_{ijr} denotes the kick given to i by j on the rth revolution. Let us calculate $\varepsilon_i^2(\tau_2) - \varepsilon_i^2(0)$ from this for the TT cooling case:

$$\begin{aligned}
\varepsilon_i^2(\tau_2) - \varepsilon_i^2(0) &= -2g\varepsilon_i \Sigma\, \Sigma\, \delta_{ijr} + g^2\, (\Sigma\, \Sigma\, \delta_{ijr})^2 \\
&= -2gMb\varepsilon_i^2 \qquad + g^2M^2 \Big(\sum_{j\ \varepsilon_j - \varepsilon_i < \sigma_0/N_s} \delta_{ij1}\Big)^2 \qquad \text{(A3)} \\
&= -2gNb\varepsilon_i \qquad + g^2M^2\, \Sigma\, \delta_{ij1}^{\;2} \\
&= -2gMb\varepsilon_i^2 \qquad + g^2M^2\, \frac{NE_0\ \tau_{PU}}{M\ T_0\, \eta'_{TT}}\ \psi(\varepsilon)\ \overline{\delta_{ij}^{\;2}}.
\end{aligned}$$

In the second line we use the linearization:

$$\delta_{ijr} \approx e\mathcal{E}_0 \tau_i \int_{-\infty}^{\infty} h''(t)\, f(t)dt$$

$$\Rightarrow \qquad b = e\mathcal{E}_0 \frac{\eta' T}{E_0} \int_{-\infty}^{\infty} h''(t) f(t)dt. \tag{A4}$$

The sum in the diffusion term is reduced by noting that only a very narrow (M dependent) band of revolution frequencies near ε_i will effect particle i after averaging M turns. This band contains $(NE_0\tau_{PU}/MT_0\eta'_{TT})\psi(E)$ particles. However, this subset is randomly distributed azimuthally, an average properly taken into account in $\overline{\delta^2}_{ij}$. Note that $\overline{\delta}_{ij}^2$ is independent of i (and ε_i). Integrating (A3) gives the second moment rate:

$$\dot{\sigma}_\varepsilon^{\,2} = -2(G/T_0)\sigma_\varepsilon^{\,2} + T_0(G/T_0)^2 N\alpha\overline{\delta}_{ij}^{\,2}\int_{-\infty}^{\infty} \psi^2(\varepsilon)d\varepsilon \tag{A5}$$

with $\qquad \alpha = E_0\tau_{PU}/\eta'_{TT}T_0$ (dimensionless).

It is tempting to immediately make the identification [via comparison with the second moment of (17)]

$$g = (G_0/T\ )$$
$$\kappa_{TT} = \frac{1}{2} N\alpha\overline{\delta^2_{ij}}.$$

This result is correct, but fortuitous. There is an inconsistency in the derivation (A3) since the diffusion term is **second** order in the interaction (g^2) but I have implicitly kept only first-order (g) terms in the friction term. It turns out[19] that part of the second-order contribution from $\sum_r \sum_j \delta_{ijr}$ can be identified with the diffusion coefficient and incorporated into the "diffusion term".[20] This manipulation changes the form of the Fokker-Planck equation but not the coefficient κ_{TT}.

The other neglected higher order part of $\Sigma\ \Sigma\ \delta_{ijr}$ has to do with the "beam feedback."[4] In order to heuristically analyze all these contributions we may imagine that the electric field kick seen by the ith particle is made up of:[21]

$$\mathcal{E}_i \equiv \mathcal{E}_s + \mathcal{E}_{pol} + \mathcal{E}_f, \tag{A6}$$

where $\mathcal{E}_s$ = single particle coherent influence $\propto$ g

$\mathcal{E}_{pol}$ = polarization field ith induces on the other particles $\propto g^2$

$\mathcal{E}_f$ = the fluctuation field of the other particles $\propto$ g.

$\mathcal{E}_s$ alone gives the first term in (A5). $\mathcal{E}_{pol}$ generates a beam feedback correction, which may be written as a gain diminution in the friction term.[22] In higher orders one encounters products of the above fields. In second order the "$\mathcal{E}_f \times \mathcal{E}_f$" term contributes to the diffusion term ($\propto g^2$), while $\mathcal{E}_{pol} \times \mathcal{E}_o$ is the g^3 correction to the **diffusion** term due to beam feedback. An illuminating discussion of the beam feedback (originally given by Sacherer[2]) is in Ref. 4, while a canonical plasma-physics approach is presented in Ref. 23 The situation is summarized in Fig. 7. Polarization induced on the beam by **any** particle will smear out by dispersion if the K → PU dispersion is great enough. One expects this to be the case with TT cooling.

The remaining task is to evaluate $\overline{\delta^2_{ijr}}$ (i independent and averaged over j):

$$\delta_{\substack{ij \\ j=1}} = e\,\mathcal{E}_0 \int_{-T/2}^{+T/2} h'(t - \tau_i)f(t + \Delta t_{ij})dt \tag{A7}$$

$$= e\,\mathcal{E}_0 \mathrm{Re} \int_{-\infty}^{\infty} \omega h(\omega)f(\omega)e^{-i\omega(\Delta t_{ij}-\tau_i)}d\omega,$$

where $h(\omega)$, $f(\omega)$ are the Fourier transforms of $h(t)$, $f(t)$. I simplify the mathematics by assuming that $h(\omega) = f(\omega)$ (similar PU and K electrical response), and that $h(\omega)$ is rectangular (see Fig. 5c). I now use the fact that $(\Delta t_{ij} - \tau_i)$ is random in a sum over j (random azimuthal distribution) to convert the sum over j in $\overline{\delta^2_{ij}}$ to an integral over $x \equiv \Delta t_{ij} - \tau_{ij}$.

$$\overline{\delta^2_{ij}} = e^2 \mathcal{E}^2_0 \iint \omega\omega' h^2(\omega)h^2(\omega')e^{-i(\omega+\omega')x}d\omega d\omega'$$

$$\Rightarrow \quad \int_{-T_0/2}^{T_0/2} \delta^2 \, dx = e^2 \mathcal{E}^2 2\pi \iint \omega\omega' h^2(\omega)h^2(\omega')\delta(\omega + \omega')d\omega d\omega' \tag{A8}$$

$$= 4\pi e^2 \mathcal{E}^2_0 \int_0^{\infty} \omega^2 h^4(\omega)d\omega.$$

However, there is a normalization equal to T_0/N for the x integral. Therefore:

$$\overline{\delta^2_{ij}} = \frac{4\pi}{T_0} e^2 \; {}^2_0 \int_0^{\infty} \mathcal{E} \omega^2 h^4(\omega)d\omega \tag{A9}$$

$$\propto \frac{1}{T_0} e^2 \mathcal{E}^2 \; (2W_{eff})^{-1} \frac{1}{3},$$

where the last line uses the rectangular approximation for $h(\omega)$.

For FT cooling the evaluation, starting with (A3), of δ_ε^2 is identical except that $\overline{\delta_{ij}^2}$ is now ε_i dependent. Also α will be a different constant since it is η dependent and factors $\sim \sqrt{2}$ as in (11) will appear. Using the same linearization as in (A4) we have

$$\overline{\delta_{ij}^2}\ (\mathrm{FT}) = \frac{4\pi}{T_0}\ e^2 \mathcal{E}_0^2\ \tau_i{}^2 \int_0^\infty \omega^4 h^4(\omega)d\omega$$

$$\propto \frac{1}{T_0}\ e^2 \mathcal{E}_0^2\ \frac{\tau_i{}^2}{(2W_{eff})^2}\ (2W_{eff})^{-1}\ \frac{1}{5}, \tag{A10}$$

where the proportionality constant is the same as in (A9). The factor $\tau_i^2/(2W)^2 \equiv \varepsilon_i^2/\varepsilon_i^{*2}$, which one may absorb into the ε_i sum. The ratio of diffusion coefficients is thus

$$\frac{\text{Diffusion TT}}{\text{Diffusion FT}} = 2\ \frac{5}{3}\ \{\int\ (\varepsilon/\varepsilon^*)^2\psi^2(\varepsilon)\}^{-1} \tag{A11}$$

Then, translated into the definition of κ defined by (1) and (17) we obtain (13). Notice that the factor $(2T_0W_{eff})^{-1}$ in both (A9) and (A10) make the diffusion terms proportional to N_S, as expected.

Finally, what is called "mixing factor" in (13) is exactly proportional to η; that is, $MF_{TT}/MF_{FT} = \eta_\pi^*/\eta_{FT}^* = \eta'/\eta$.

APPENDIX II

Since the instantaneous rate (2) is constrained by $\int \varepsilon^2 \psi^2 d\varepsilon / \sigma_\varepsilon^2$ we can seek an upper bound to (2) by finding a minimum for $\Phi(\psi) = \int \varepsilon^2 \psi^2 d\varepsilon / \sigma_\varepsilon^2$. For σ_ε^2 fixed we have two constraints on ψ_{min}: 1) it is normalized and 2) from (7) we see that,

$$\dot{\psi}(\varepsilon=0) = g\ \psi(\varepsilon=0). \tag{A12}$$

(Of course this experimental central density growth stops at some level due to circuit imperfections we ignore.) Unfortunately one can still choose distributions ψ_{min} such that $\Phi \rightarrow 0$ $t \rightarrow \infty$. This is reflected in the fact that the asymptotic solution to (1) is singular at the origin. Qualitatively this mathematical problem leads to unphysical distributions with large tails contributing to σ_ε^2 and sharp central spikes contributing to $\int \varepsilon^2 \psi^2 d\varepsilon$. Examination of (7), however, shows the damping to be strongest out in the distribution tail. Therefore we insist on limiting the distribution width as well as σ_ε.

The minimization problem is now considerably better behaved. A reasonable approximation for ψ_{min} is then a Gaussian. This choice gives

$$\Phi(\text{Gaussian}) = \frac{1}{\sigma_\varepsilon} \ \frac{1}{4\sqrt{\pi}} = \psi(\varepsilon=0)\ \frac{\sqrt{2}}{4}. \tag{A13}$$

For fixed g_0 we find a solution for ΔT (F)[substituting (A13) into Eq.(7)]

$$\Delta T = \frac{1}{g_0} \operatorname{Ln} \frac{F}{2-F}. \tag{A14}$$

Further it is easy to see that the rate is optimized by setting $g_{max}(t) \equiv g_0 \, \sigma_\varepsilon^{(t)} / \sigma_\varepsilon(t)$. This gives a solution

$$\Delta T = \frac{1}{g_0} (F-1). \tag{A15}$$

These bounds have been obtained using some rather drastic assumptions about the distribution shape. Finally one must study the exact nonlinear evolution of (7). However this qualitative flow is substantiated by some recent numerical computations by Crosbie[24] who has simulated the various cases treated in this paper. I compare his results with the above estimates in Fig. 8.

REFERENCES

[1] The Fermilab High-Intensity Antiproton Source Design Report, Oct. 1979 (incorporated in Tevatron Phase I Project Design Report, Feb. 1980).

[2] F. Sacherer, Stochastic Cooling Theory, CERN/ISR/TH 78-11.

[3] W. Kells, Methods for Longitudinal Cooling by Large Factors in a $\bar{p}$ Precooler Ring, Fermilab $\bar{p}$ note 43, Oct. 1979.

[4] D. Möhl et al., Theory and Technique of Stochastic Cooling, CERN/PS/AA 79-23, July 1979.

[5] The naive derivation of rate equations for the distribution r.m.s. by summing the squares of kicks for each revolution give incorrect diffusion coefficients. The integral $\int \varepsilon^2 \psi^2 d\varepsilon$ appearing in Eq. (2) is not reproduced. The essential nonlinearity is missed in this approach.

[6] D. Möhl, Stochastic Cooling, CERN/PS/DL 78-25.

[7] Most of the early stochastic cooling papers invoke such a derivation. See the first section of Refs. 4 and 6.

[8] F. Sacherer, **ibid.**, p.10, equation 24.

[9] R. B. Palmer, see reference cited and discussion in Ref. 4 above.

[10] H. G. Hereward, The Elementary Theory of Landau Damping, CERN 65-20, 1965. See also Ref. 7 above.

[11] S. Ichimaru, **Basic Principles of Plasma Physics** (W. Benjamin, Reading, Mass, 1973). See Chapters 2 and 10.

[12] Design Study of a Proton-Antiproton Colliding Beam Facility, CERN/PS/AA 78-3.

[13]A survey of available commercial power amplifiers in the UHF range shows power × bandwidth to be limited to $\lesssim$ 4000 MHz watt for regions of flat gain and phase delay.

[14]W. Kells, Waveguide Longitudinal Pick-Up, Fermilab $\bar{p}$ note 56, 1979.

[15]P. M. McIntyre, A Filter Technique for Stochastic Cooling of Betatron Oscillations, Fermilab Internal Report TM-920, 1979.

[16]H. Herr and D. Möhl, Bunched Beam Stochastic Cooling, CERN EP note 79-34.

[17]J. Griffin (private communication).

[18]J. Bisognano, Kinetic Equations For Longitudinal Stochastic Cooling, LBL note BECON-11, 1979.

[19]S. Ichimaru, **ibid.**, p 240-241.

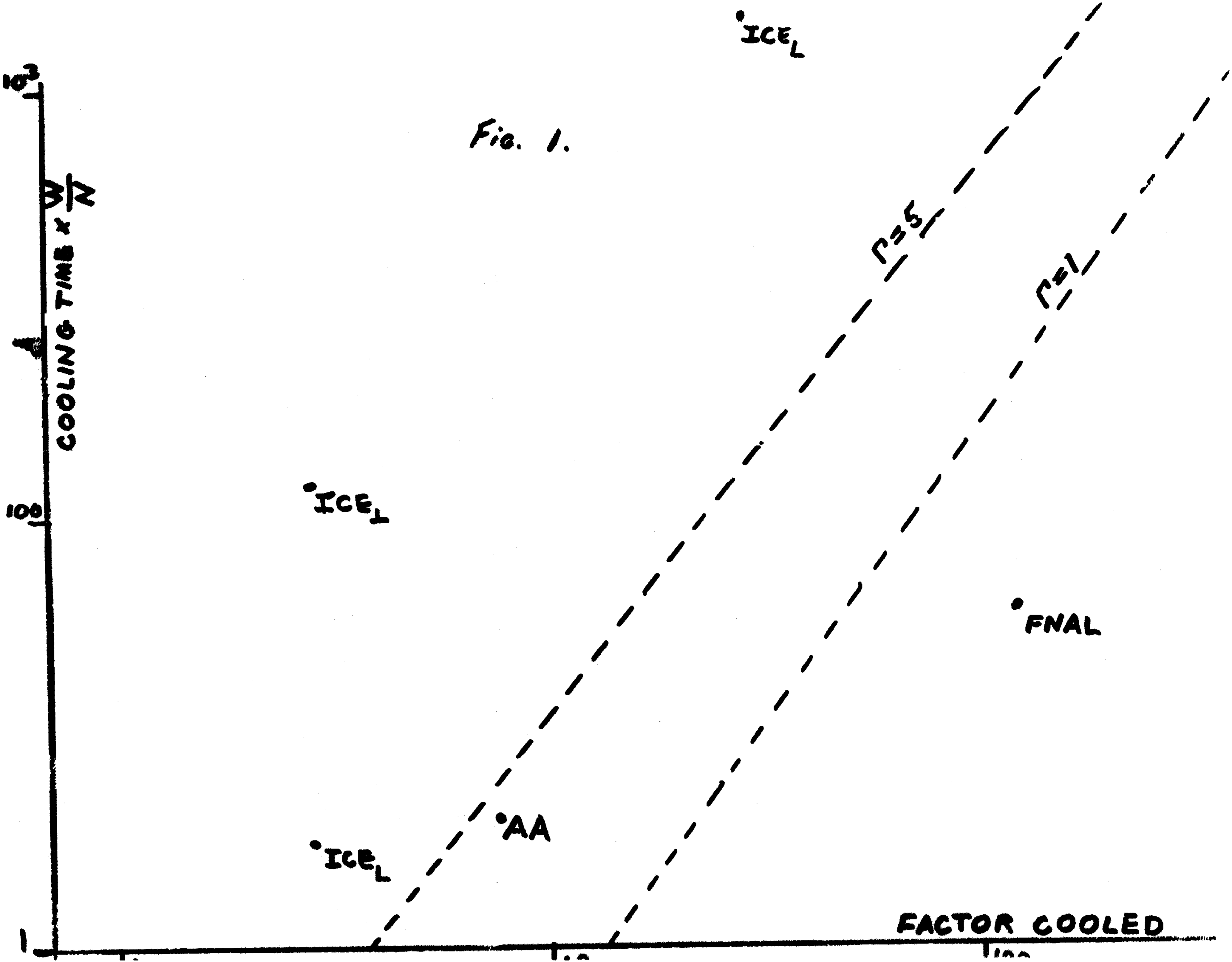
[20]See Ref. 4 p.

[21]S. Ichimaru, **ibid.**, p. 239.

[22]One efficient way of combining the higher order feedback terms is developed in Ref. 4, p

[23]J. Bisognano, Vertical Transverse Stochastic Cooling, LBL note BECON-10, 1979.

[24]E. Crosbie, A Program for STUB FILTER Cooling in the Time Domain, Fermilab $\bar{p}$ note 52, 1979.



10^3
100
1
COOLING TIME $\times \frac{W}{N}$
FACTOR COOLED
ICE_L
ICE_L
ICE_L
AA
FNAL
$\Gamma=5$
$\Gamma=1$

Fig. 1.

FIGURE 2

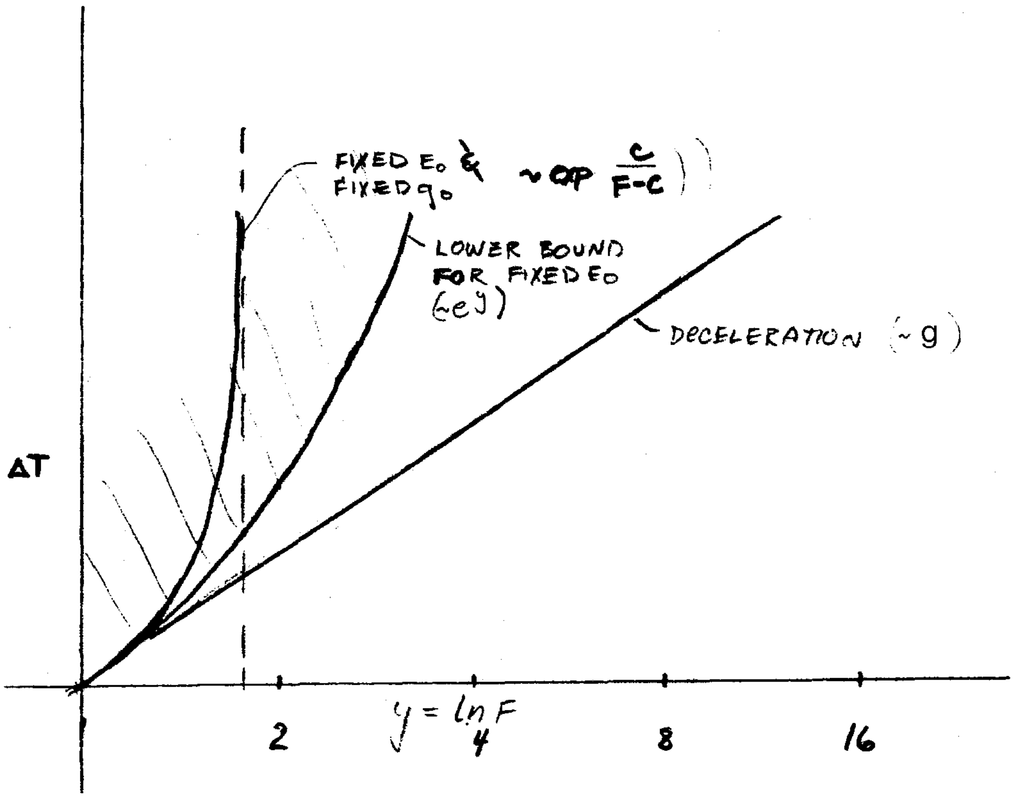
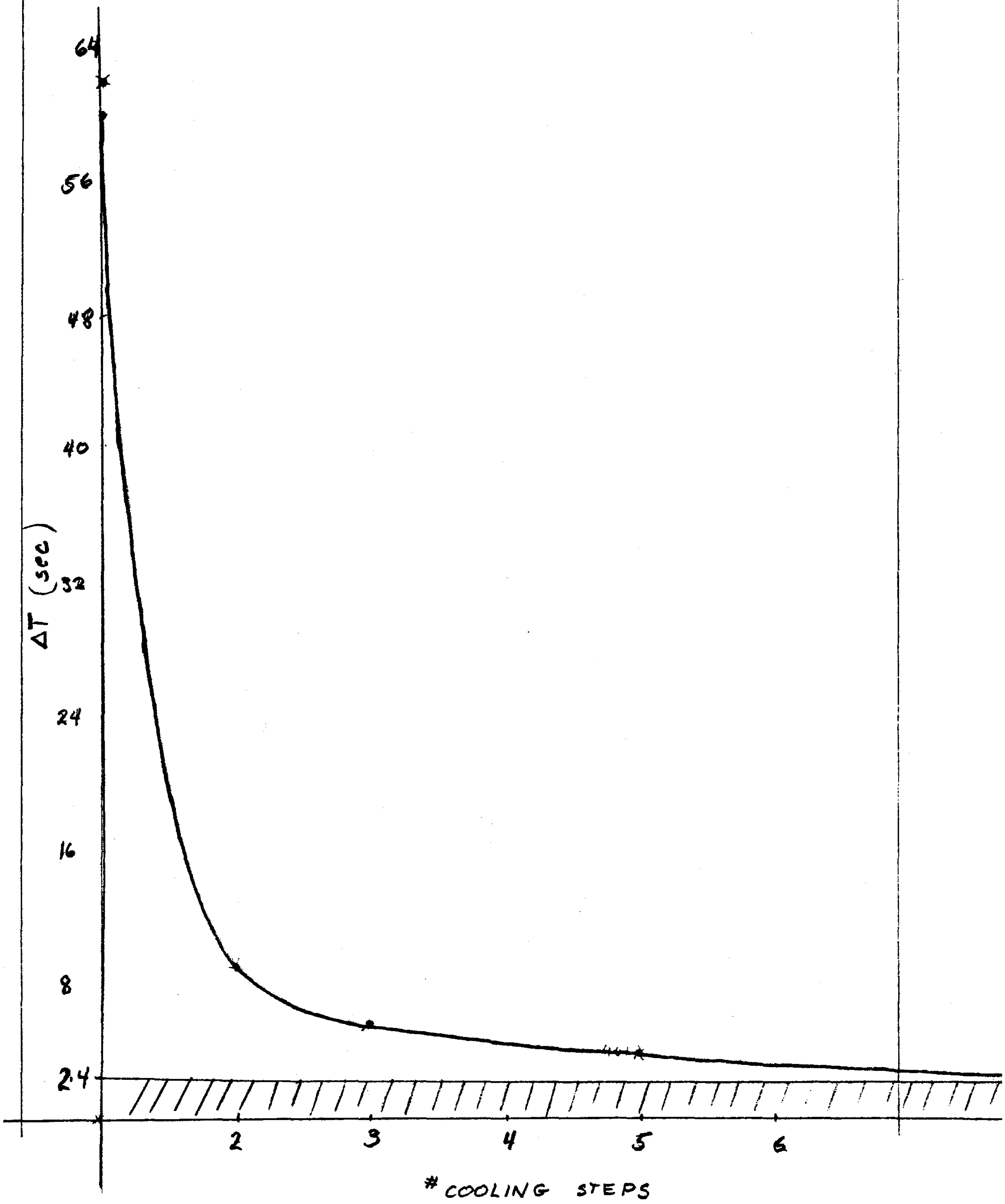
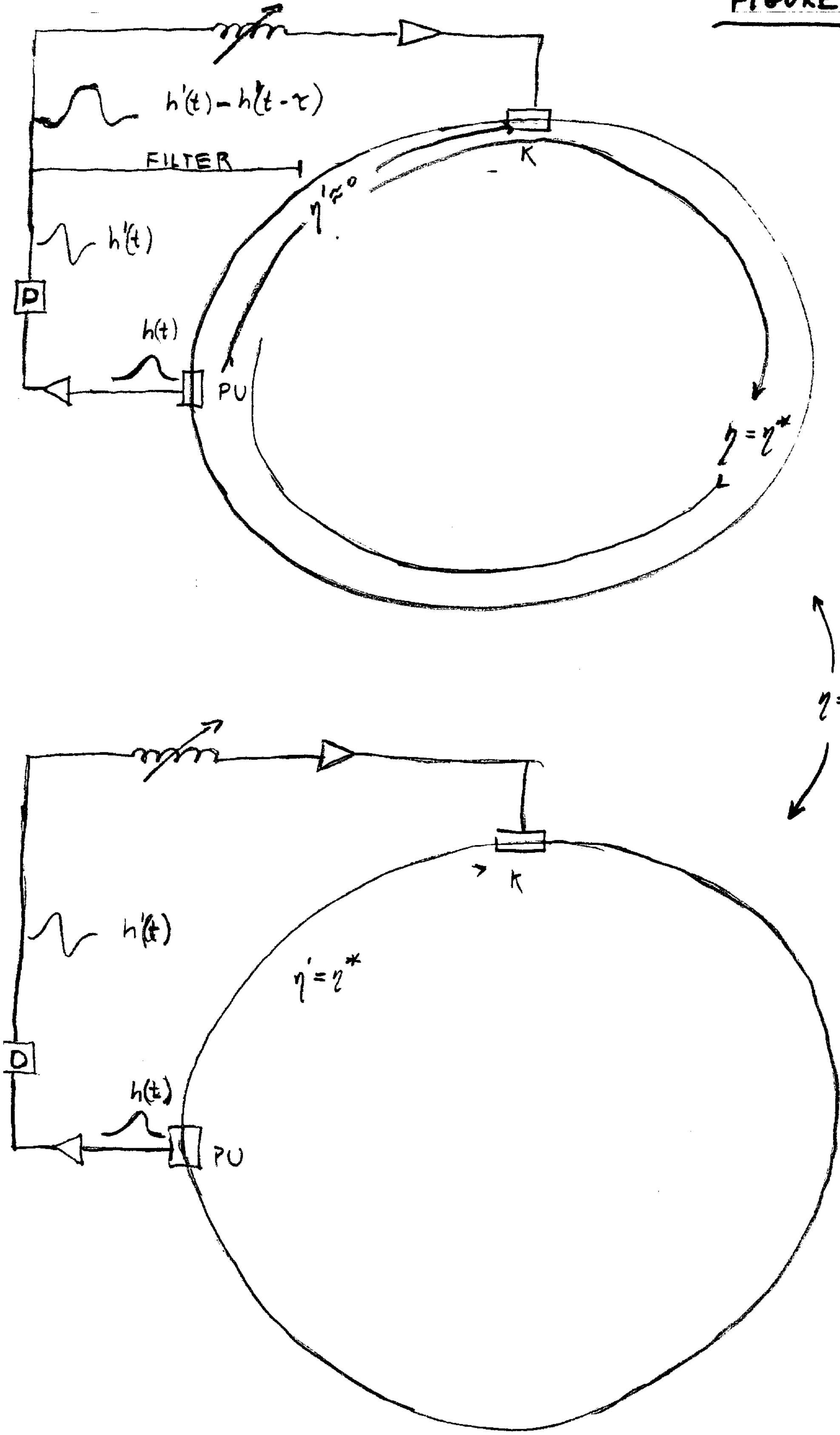

FIXED E₀ & FIXED q₀ (~exp c/(F-c))
LOWER BOUND FOR FIXED E₀ (~e^y)
DECELERATION (~g)
ΔT
y = ln F
2
4
8
16

FIGURE 3



<u>FIGURE 4</u>



4a

FILTER COOLING

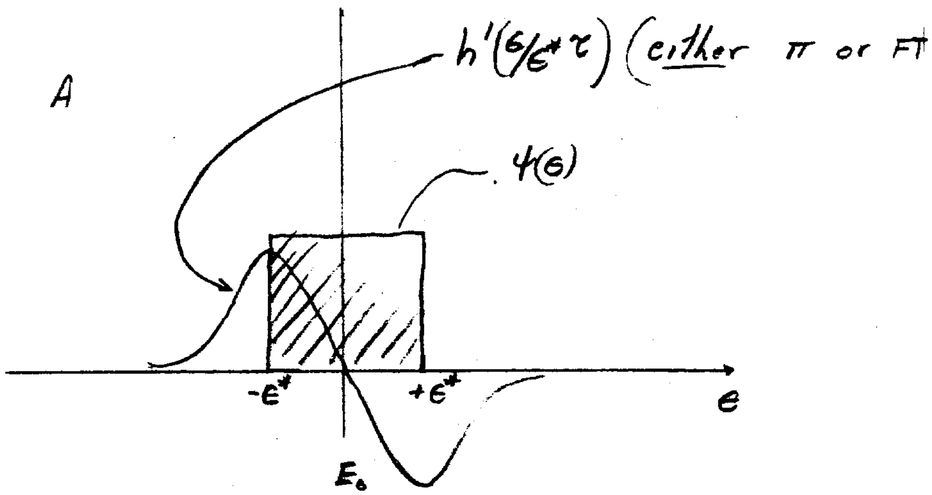
η = entire ring dispersion

4b

TRANSIT TIME COOLING

FIGURE 5

A



τ ≈ 1/2W_eff

(B)

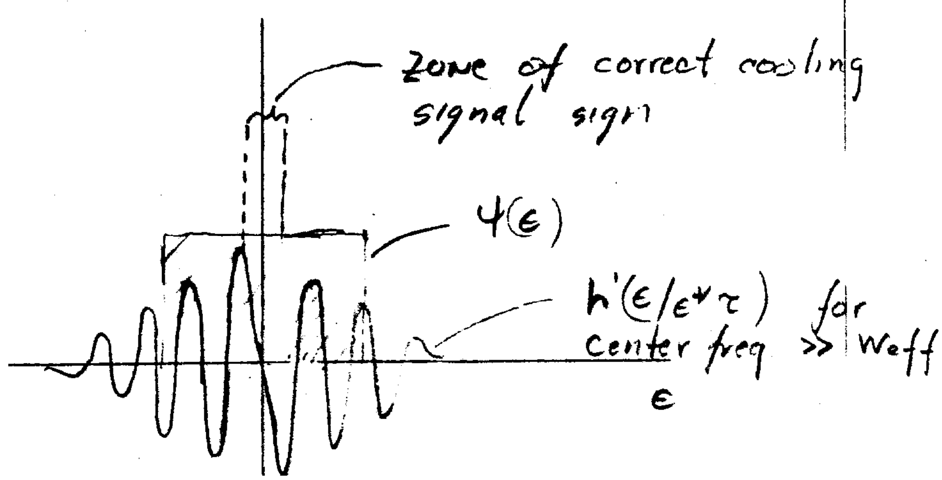


FIGURE 6

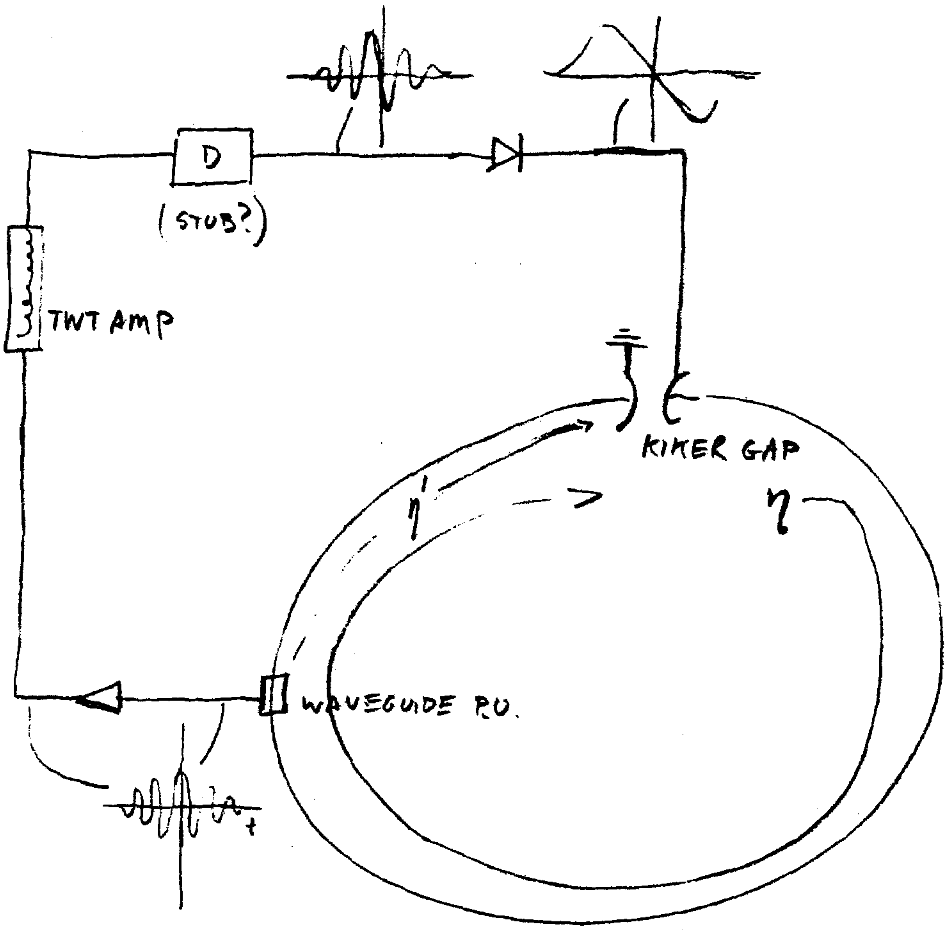
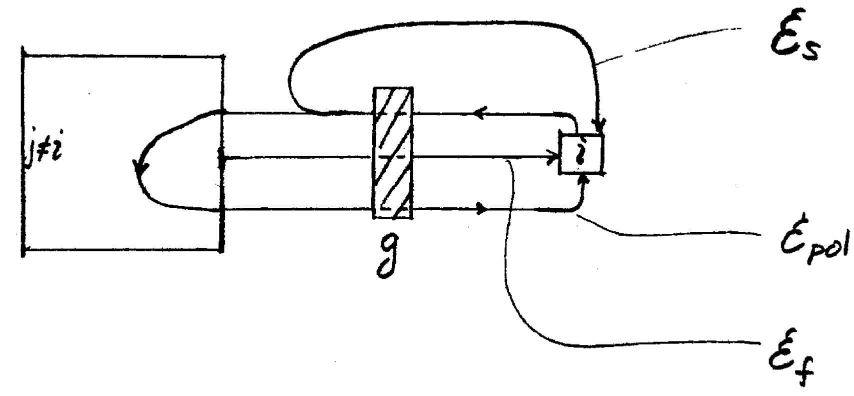


FIGURE 7



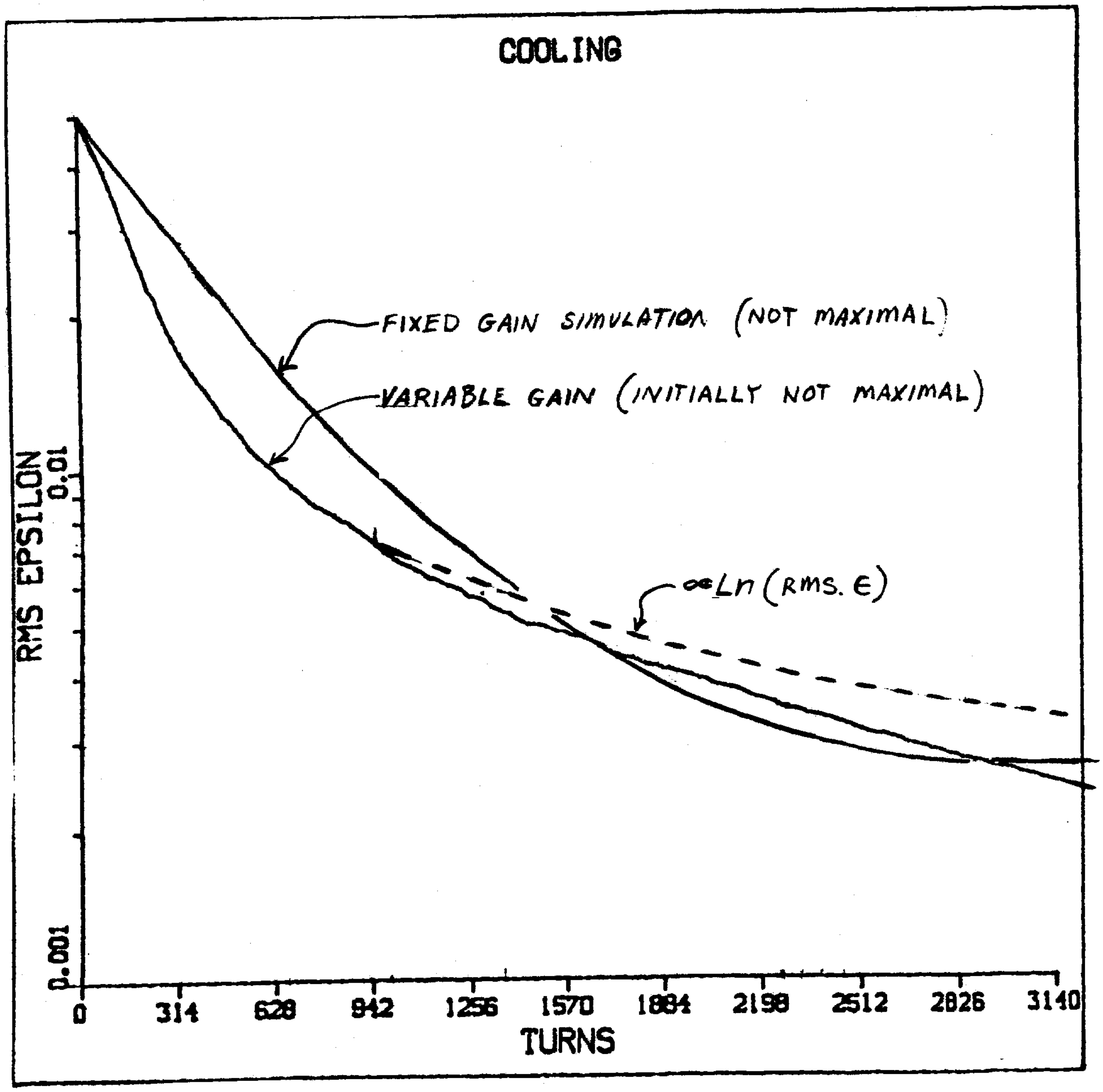
F TERM

1^{ST} ORDER:	$\mathcal{E}_s$	—	CERN FRICTION TERM	g
	$\langle \mathcal{E}_{pol} \rangle$	—	BEAM FEEDBACK CORRECTION TO	g^2
	$\langle \mathcal{E}_f \rangle$	—	$\sim g \, 1/\sqrt{N} \sim 0$	g
2^D ORDER:	$\langle \mathcal{E}_s \cdot \partial \mathcal{E}_s / \partial r \rangle$	—	$\propto 1/N \sim 0$	g^2
	$\langle \mathcal{E}_f \cdot \partial \mathcal{E}_f / \partial r \rangle$	—	TO BE LUMPED IN D TERM	g^2
	$\langle \mathcal{E}_{pol} \cdot \partial \mathcal{E}_f / \partial r \rangle$	—	BEAM FEEDBACK CORRECTION TO	g^3
	$\langle \mathcal{E}_s \cdot \partial \mathcal{E}_{pol} / \partial r \rangle$	—	$\sim 1/N \sim 0$ relative to	g^3
	$\langle \mathcal{E}_s \cdot \partial \mathcal{E}_f / \partial r \rangle$	—	~ 0	g^2
	$\langle \mathcal{E}_{pol} \cdot \partial \mathcal{E}_{pol} / \partial r \rangle$	—	$O(g^4)$	g^4

D TERM

1^{ST} ORDER:	$\langle \mathcal{E}_f \cdot \mathcal{E}_f \rangle$	—	USUAL DIFFUSION TERM	g^2
	$\langle \mathcal{E}_f \cdot \mathcal{E}_s \rangle$	—	~ 0	
	$\langle \mathcal{E}_f \cdot \mathcal{E}_{pol} \rangle$	—	Beam FEEDBACK TO CORRECTION TO	g^3
	$\langle \mathcal{E}_{pol} \cdot \mathcal{E}_{pol} \rangle$	—	$O(g^4)$	

FIGURE 8


COOLING
FIXED GAIN SIMULATION (NOT MAXIMAL)
VARIABLE GAIN (INITIALLY NOT MAXIMAL)
∝Ln (RMS. ∈)
RMS EPSILON
0.01
0.001
0
314
628
942
1256
1570
1884
2198
2512
2826
3140
TURNS